



## Simulation de génotypes intégrant du DL avec des DAG

---

**François GUILLAUME**

31 mars 2011

INRA-GABI, Institut de l'Élevage

# Comment simule-t-on du DL ?

## La méthode classique

Meuwissen Hayes et Goddard (2001), population 1000 générations jusqu'à l'équilibre mutation dérive.

Simule-t-on quelque chose qui ressemble à ce qu'on constate dans les faits ?

## Comment simule-t-on du DL ?

### Une méthode rustre

Durant ma thèse (2007), écriture de genot, un programme de transmission de phases fondatrices au sein d'un pedigree réel (+ script shell intégrant genot)

- On stocke des phases des fondateurs de la population (phases considérées "vraies")
- On les réattribue aléatoirement à la population de fondateurs (avec répétitions)
- Transmission dans le pedigree + recombinaisons
- On exclu un SNP satisfaisant  $Var_{QTL} = (2pq)a^2$  (servira à simuler le phénotype)

Dispose-t-on de suffisamment de données pour avoir un échantillonnage correct ?

## Évolutions récentes

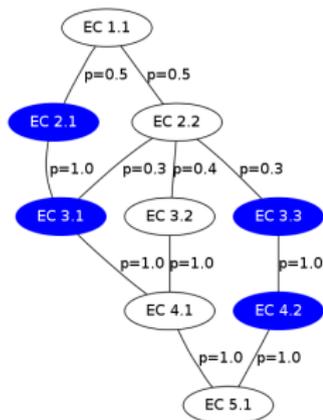
- Accumulation de données
- Méthodes de reconstruction de phases plus raffinées (HMM)
- Imputation

Note : Imputer c'est prédire l'allèle **le plus probable** à un marqueur sachant les allèles présents aux autres marqueurs  
L'idée serait de se servir du modèle de description du DL utilisé en imputation/reconstruction de phases pour simuler des données de typages

FastPhase propose cette option =>cf Présentation de Bertrand septembre dernier

## Idée

Comme pour FastPhase, on se base sur un modèle de description de la structure du DL (DAG) pour simuler des phases présentant un DL réaliste. Pour chaque nouvel état caché on tire une valeur d'une loi uniforme sur  $[0,1]$   $\Rightarrow$  probabilité de transition vers un état au marqueur suivant.



## Après....

- Une sous-routine de lecture et stockage du DAG a été écrite
- Intégration (+modification de genot) dans genot commencé
- Reste à tester / améliorer